



Brainomics: A management system for exploring and merging heterogeneous brain mapping data

Vincent Michel, Yannick Schwartz, Philippe Pinel, Olivier Cayrol, Antonio Moreno, Jean-Baptiste Poline, Vincent Frouin, Dimitri Papadopoulos Orfanos

► To cite this version:

Vincent Michel, Yannick Schwartz, Philippe Pinel, Olivier Cayrol, Antonio Moreno, et al.. Brainomics: A management system for exploring and merging heterogeneous brain mapping data. OHBM 2013 19th Annual Meeting of the Organization for Human Brain Mapping, Jun 2013, Seattle, United States. <cea-00904768>

HAL Id: cea-00904768

<https://hal-cea.archives-ouvertes.fr/cea-00904768>

Submitted on 15 Nov 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Brainomics - A management system for exploring and merging heterogeneous brain mapping data

V. Michel¹, Y. Schwartz^{2,4}, P. Pinel³, O. Cayrol¹, A. Moreno³,
JB. Poline⁴, V. Frouin⁴, D. Papadopoulos Orfanos⁴

¹Logilab, Paris, France, ²INRIA Saclay Parietal Team, Paris, France,

³Unicog, INSERM-CEA, Neurospin, Gif-sur-Yvette, France, ⁴CEA, Neurospin, Gif-sur-Yvette, France



Introduction

- Number of large datasets for brain mapping have been released [1, 2].
- Neuroimaging datasets more routinely include clinical data or genetics data.
- Exploitation requires
 - An efficient organization to integrate all the measures
 - An easy access to the relevant information.
- Neuroimaging [3] and genomics[4] databases are dedicated to their own field of research.

<http://www.brainomics.net/demo/>

- Brings together **brain imaging and genetics data**.
- Relies on a **high-level query language (RQL)**.
- Solution based on **CubicWeb**, a semantic framework.
- Visualizing / exporting data** in several formats.

question	text	score	date
ados_module	Autism Diagnostic Observation Schedule		2010-03-19
ados_social	Social Total Subscore of the Classic ADOS	9.0	2010-03-19
ados_stereo_behav	Stereotyped Behaviors and Restricted Interests Total Subscore of the Classic ADOS	4.0	2010-03-19
ados_rsrch_reliable	Was ADOS scored and administered by research reliable personnel?	0.0	2010-03-19
ados_gotham_soc_affect	Social Affect Total Subscore for Gotham Algorithm of the ADOS	4.0	2010-03-19
ados_gotham_rrb	Restrictive and Repetitive Behaviors Total Subscore for Gotham Algorithm of the ADOS	7.0	2010-03-19
ados_gotham_total	Social Affect Total + Restrictive and Repetitive Behaviors Total	16.0	2010-03-19
ados_gotham_severity	Individually Calibrated Severity Score for Gotham Algorithm of the ADOS	6.0	2010-03-19



<http://www.cubicweb.org/>

- ✓ **Data management framework**, 10 years of industrial uses (e.g. [5]).
- ✓ Well established core technologies: **SQL, Python, HTML5, Javascript**.
- ✓ Licensed under **LGPL** since 2008.
- ✓ Used in production environments since 2005.
- ✓ Fine-grained **security system** coupled to the data model definition.
- ✓ **Migration mechanisms** controls model version / ensures data integrity.

Data model

- ✓ Described in Python, using reusable modules called "cubes".
- ✓ Modelisation of Scans, Questionnaires, Genomics results, Behavioral results, Subjects and Studies information.
- ✓ Data model optimized for large volumes (> 2000 subjects).
- ✓ Tested with several publicly available datasets [1, 2].

Query using RQL

- ✓ Similar to the W3C's SPARQL [6].
- ✓ Supports the basic operations (selection, insertion, etc.).
- ✓ Subquerying, ordering, counting, ...

Query all the female subjects of the database, with an age greater than 30
Any S WHERE S is Subject, S age > 30, S gender "female"

Query all the Cmap scans of subjects with an age greater than 25, and that have a score greater than 4.0 for the "algebre" question
Any SA WHERE S is Subject, S age > 25, X is QuestionnaireRun, X concerns S, A is Answer, A questionnaire_run X, A question Q, Q text "algebre", A value > 4, SA is Scan, SA concerns S, SA type "c map"

Views

- ✓ Each query result can be seen using different views.
- ✓ HTML pages, ZIP files, spreadsheets, XCEDE XML, ...
- ✓ May include processing (stat. maps computed on the fly).

Conclusion

- ✓ **Open source solution to manage brain imaging datasets and associated meta data.**
- ✓ **Powerful querying and reporting tool, customized for emerging imaging-genetics field.**

Contact: brainomics@logilab.fr

This work was supported by grants from the French National Research Agency (ANR GENIM; ANR-10-BLAN-0128) and (ANR IA BRAINOMICS; ANR-10-BINF-04).

[1] <http://openfmri.org/data-sets>

[2] fcon_1000.projects.nitrc.org/indi/abide/

[3] Olsen, M.D. (2007). The extensible neuroimaging archive toolkit. Neuroinformatics 5, 11-33

[4] Vallon-Christersson, J. (2009). 'BASE - 2nd generation software for microarray data management and analysis'

[5] Publishing bibliographic records on the Web of data: opportunities for the BnF (French national Library). ESWC 2013

[6] <http://www.w3.org/TR/rdf-sparql-query/>